Claims

What is claimed is:

1. A method of processing a database query, the query including one or more expressions, the method including:

resolving columns in one or more of the expressions;

performing expression optimization on one or more of the expressions;

performing further query optimization; and

where the expression optimization is performed before further query optimization.

2. The method of claim 1, where each expression includes one or more sub-expressions, and where the expression optimization includes, for each expression:

(1) if the expression has a form selected from the group consisting of "SE+0," "SE*1," and "SE/1," where SE is a sub-expression, then reducing the expression to SE;

(2) if the expression has a form selected from the group consisting of "SE*0," "0/SE," and "0 MOD SE," where SE is a non-nullable sub-expression, then reducing the expression to 0; and

(3) if the expression is of the form F(C), where F is a function and C is a constant and F(C) returns the return value, then reducing the expression to a return value.

3. The method of claim 2, where one or more of the sub-expressions include sub-expressions, the method including:

(4) for each sub-expression that includes a sub-expression, simplifying the sub-expression using (1)-(3).

4. The method of claim 2, where SE is nullable if it includes a nullable column.

5. The method of claim 2, where SE is nullable if it belongs to an inner table of an outer join.

6. The method of claim 1, where the query is represented by a tree, including one or more nodes.

7. The method of claim 1, where the query includes an assignment list clause and where one or more of the expressions are in the assignment list clause.

8. The method of claim 1, where the query includes a WHERE clause, and where one or more of the expressions are in the WHERE clause.

9. The method of claim 1, where further query optimization includes:
determining a satisfiability of the database query.

10. The method of claim 1, where further query optimization includes:
determining a transitive closure of the database query.

11. The method of claim 1, where further query optimization includes:
determining one or more plans for executing the query.

12. The method of claim 11, where one of the one or more plans includes:
scanning a table to locate rows that satisfy one or more conditions; and
summing one or more columns in the rows that satisfy the one or more conditions.

13. The method of claim 1, where further query optimization includes:
selecting an optimal plan from executing the database query.

14. The method of claim 1, where further query optimization includes two or more optimizations selected from the group consisting of:
determining a satisfiability of the database query;
determining a transitive closure of the database query;
determining one or more plans for executing the query; and
selecting an optimal plan from executing the database query.

15. A computer program, stored on a tangible storage medium, for use in processing a database query, the query including one or more expressions, the computer program including executable instructions that cause a computer to:

resolve columns in one or more of the expressions;

perform expression optimization on one or more of the expressions;

perform further query optimization; and

where the expression optimization is performed before further query optimization.

16. The computer program of claim 16, where each expression includes one or more sub-expressions, and where the computer program including executable instructions that cause a computer to, for each expression:

(1) determine if the expression has a form selected from the group consisting of "SE+0," "SE*1," and "SE/1," where SE is a sub-expression, and if so, then reduce the expression to SE;

(2) determine if the expression has a form selected from the group consisting of "SE*0," "0/SE," and "0 MOD SE," where SE is a non-nullable sub-expression, and if so, then reduce the expression to 0; and

(3) determine if the expression is of the form F(C), where F is a function and C is a constant and F(C) returns the return value, and if so, reduce the expression to a return value.

17. The computer program of claim 16, where one or more of the sub-expressions include sub-expressions, the computer program including executable instructions that cause a computer to:

(4) for each sub-expression that includes a sub-expression, simplify the sub-expression using (1)-(3).

18. The computer program of claim 16, where SE is nullable if it includes a nullable column.

19. The computer program of claim 16, where SE is nullable if it belongs to an inner table of an outer join.

20. The computer program of claim 15, where the query is represented by a tree, including one or more nodes.

21. The computer program of claim 15, where the query includes an assignment list clause and where one or more of the expressions are in the assignment list clause.

22. The computer program of claim 15, where the query includes a WHERE clause, and where one or more of the expressions are in the WHERE clause.

23. The computer program of claim 15, where further query optimization includes:
determining a satisfiability of the database query.

24. The computer program of claim 15, where further query optimization includes:
determining a transitive closure of the database query.

25. The computer program of claim 15, where further query optimization includes:
determining one or more plans for executing the query.

26. The computer program of claim 25, where one of the one or more plans includes:
scanning a table to locate rows that satisfy one or more conditions; and
summing one or more columns in the rows that satisfy the one or more conditions.

27. The computer program of claim 15, where further query optimization includes:
selecting an optimal plan from executing the database query.

28. The computer program of claim 15, where further query optimization includes two or more optimizations selected from the group consisting of:
determining a satisfiability of the database query;
determining a transitive closure of the database query;
determining one or more plans for executing the query; and
selecting an optimal plan from executing the database query.

29. A database system including:

a massively parallel processing system including:

one or more nodes;

a plurality of CPUs, each of the one or more nodes providing access to one or more CPUs;

a plurality of data storage facilities each of the one or more CPUs providing access to one or more data storage facilities;

a process for execution on the massively parallel processing system for processing one or more database queries, each query including one or more expressions, the process including:

resolving columns in one or more of the expressions;

performing expression optimization on one or more of the expressions;

performing further query optimization; and

where the expression optimization is performed before the further query optimization.

30. The database system of claim 29, where each expression includes one or more sub-expressions, and where the expression optimization includes, for each expression:

(1) if the expression has a form selected from the group consisting of "SE+0," "SE*1," and "SE/1," where SE is a sub-expression, then reducing the expression to SE;

(2) if the expression has a form selected from the group consisting of "SE*0," "0/SE," and "0 MOD SE," where SE is a non-nullable sub-expression, then reducing the expression to 0; and

(3) if the expression is of the form F(C), where F is a function and C is a constant and F(C) returns the return value, then reducing the expression to a return value.

31. The database system of claim 30, where one or more of the sub-expressions include sub-expressions, the expression optimization including:

(4) for each sub-expression that includes a sub-expression, simplifying the sub-expression using (1)-(3).

32. The database system of claim 30, where SE is nullable if it includes a nullable column.

33. The database system of claim 30, where SE is nullable if it belongs to an inner table of an outer join.

34. The database system of claim 29, where the query is represented by a tree, including one or more nodes.

35. The database system of claim 29, where the query includes an assignment list clause and where one or more of the expressions are in the assignment list clause.

36. The database system of claim 29, where the query includes a WHERE clause, and where one or more of the expressions are in the WHERE clause.

37. The database system of claim 29, where further query optimization includes:
determining a satisfiability of the database query.

38. The database system of claim 29, where further query optimization includes:
determining a transitive closure of the database query.

39. The database system of claim 29, where further query optimization includes:
determining one or more plans for executing the query.

40. The database system of claim 39, where one of the one or more plans includes:
scanning a table to locate rows that satisfy one or more conditions; and
summing one or more columns in the rows that satisfy the one or more conditions.

41. The database system of claim 29, where further query optimization includes:
selecting an optimal plan from executing the database query.

42. The database system of claim 29, where further query optimization includes two or more optimizations selected from the group consisting of:
determining a satisfiability of the database query;
determining a transitive closure of the database query;
determining one or more plans for executing the query; and
selecting an optimal plan from executing the database query.